



OY	356	AATACACCTCTCGAACCAATGACCTGACCTGAGGAAGTGAAGGAGAGAGCTGCTGC	AAAG	415
Db	43	SerSerValIleuLeuThrIleAspLeuThrGluSerValAlaGluGluAlaIleSerAsn		62
OY	416	AAG---GCAGACCTCATCTCTCCATACCTCCGCGCTACTCTCCGACCCATGAAGGCATA		472
Db	63	LysLeuValSerSerIleValAlaIleThrIleProIleIlePheArgGlyLeuLysAlaIle		82
OY	473	ACCTGGAAACACATGAGACAGGACCGCTGGTATCCGGGCTCTGGAGACACAGATCGGTATC		532
Db	83	ThrMetGluAspProGlnArgSerLeuLeuLysLeuAlaAlaGluGlyIleHisVal		102
OY	533	TACTCTCCCTACAGCCTTATGATGATCCGCGCCGCGGACGTCACACACATGGTGGCTAA		592
Db	103	TyrSerProHisThrIleValAlaAspValAlaValAspGlyValAsnAspTrpLeuAlaGln		122
OY	593	GGGCTTGGAGCTTGATACCTCCAGGCGCATCATCTTCCAAAGCTCCCAACTACCTTACA		652
Db	123	GlyIleAlaGly-----		126
OY	653	GAGGAAACCAACCGAGTGAATTTCACGTTAACTACACCCAAAGACCTGGACAAAGTCATG		712
Db	127	--GlyAspAsnAsnIleLys--SerValValProThrGln-----		138
OY	713	TCTGCAGTGAAGGAATTGACGGGTGTTCTGCACTCTTTTCTGCTAGACATGTAAT		772
Db	138	-----		138
OY	773	GAGGAACAACACGGATTAATTCGTAAATTGACTCAGAAGCGTTTGATGACAGTGTAGAT		832
Db	139	-----GlnAsnSerValMetAlaGluAlaGln-----		147
OY	833	TTTCTTCCCGGAACAACAACCTTATATCAGAAGACGGAAATTCTGCTACTGGAGAAGCT		892
Db	147	-----		147
OY	893	TTTGCTTTCATACACTGGAATGGAGCGGTATGACACACTGGATGTAATGTCTCCCTGGCA		952
Db	148	-----GlyTyrGlyArgIleCysGluLeuLysIleProThrThrLeuArg		162
OY	953	ACCATGATGATTCGACATTAATAAAGACACCTAAACATATGTCATATTGCGCTTACCCCTTGGG		1011
Db	163	GluLeuValGlnArgAlaIleGluLeuThrGlyLeuGlnTyrValGlnValCysAlaPro		182
OY	1013	GTGGGAGACACCTTGAGCTGACCTCAAGTCAAGCGCGGCGCTGTGTGGTCTGGGGAGC		1072
Db	183	AsnGly-----LeuAspSerHisIleSerLysValSerLeuCysAlaGlySerIleGly		200
OY	1073	AAGCTTTCGACAGGAGTGTGAGGCTGACCTTACCTCAGACAGTGAAGATCCCATGATAT		1132
Db	201	SerValValMetAsnThrAspAlaAspLeuTyrPheThrGlnGluLeuSerHisHisGln		220
OY	1133	ACTTTGGATGCTGCTCCCAAGGAATTAATGTCACTCTGTGAACACAGACAACTGAA		1199
Db	221	ValLeuAlaAlaMetAlaIleGlyIleSerValIleLeuCysGlyHisSerAsnThrGln		240
OY	1193	CGAGCTTCTTCTTGACACTT-----CGAGATATGCTGCGATTC---CACTTGGAGAT		1242
Db	241	ArgGlyTyrLeuLysAspValMetCysGlnLysLeuAlaIleSerSerPheHisLysGluGly		260
OY	1244	--AAGATAAATATATATCTCATACAGACACTGACAGGACCCCTCTT		1285
Db	261	ValAspAlaAsnValIleValSerSerMetAspAlaAspProLeu		275

RESULT 2

S64243

hypothetical protein YGL221C - yeast (Saccharomyces cerevisiae)

N:Alternate names: hypothetical protein G1020

C:Species: Saccharomyces cerevisiae

C:Date: 17-May-1996 #sequence\_revision 17-May-1996 #text\_change 20-Jun-2000

R:Rieger, M.; Mueller-Auer, S.; Brueckner, M.; Schaefer, M.

[illegible]

Db 169 lleserleuGlunIleVallyAsnVallylsrIyglValleuArVAlPrCTyValGlN 168

Qy 1001 TTAGCCCTTGGGGGTGGGAGAACTTAGAGTCTCAAGTCAAAAGTC-----GTGGCCCTG 1054

Db 189 ValAlaSerLeuAlaAlaIaProSerAlaIaPyrAsnIleuysrIleYlsValAlaVal 208

Qy 1055 TGTGCTGCTTCTGGGACAGCGTCTGCAGAGGTGT-----GAGGCTGACCTTACCTC 1108

Db 209 CysAlaIySerGlySerGlyValPheIySeIleuIySeIuAspValAspLeuTyr 228

Qy 1109 ACAGGTGAGATGTCTCCATCATGATCTTGGATGCTGCTCCCAAGAAATAATATC 1168

Db 229 ThrIygluMetSerHisHISgluValleuIySrIySeIuMetGlyLysThrValIle 248

Qy 1169 CTCGTGGAACACAGCAACACTGAAGAGCGCTTCTTTCGACCTTCAGATATGCTG-- 1225

Db 249 ValCysAsnHisSerAntThrGluArgIyPheIuSeIuAspValMetLysGlyLeu 268

Qy 1226 -----CATTCCTGCTTGGAGAAATAGATTAATTTATTCATATGAGACTGCAGGAGC 1279

Db 269 GluAspIuGlunIyHis-----GluValValIValSerLysMetAspCysAsp 283

Qy 1280 CCTCTTCAGCTG 1291

Db 284 ProIeuThrVal 287

RESULT 3

A69954

conserved hypothetical protein yqfo - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 21-Jul-2000

C:Accession: A69954

R:Kunst, F.; Ogatawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Broillett, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Cater, N.M.; Chd, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardiniois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y.; M.; Ogawa, K.; Ogatawara, N.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelid, Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, A.; Kuchl, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamano, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: A69954

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:1: Molecule type: DNA

A:Residues: 1-373 <KUN>

A:Cross-references: GB:29116; GB:AL009126; MID:g2634723; PION:CAM14447.1; PID:g2634950

A:Experimental source: strain 168

C:Genetics:

A:Gene: yqfo

Alignment Scores:

Prod. NO.: 6,68e-16 Length: 373

Score: 285.50 Matches: 99

Percent Similarity: 46.84% Conservative: 79

Best Local Similarity: 26.05% Mismatches: 143

Query Match: 10.24% Indels: 59

DB: 2 Gaps: 17

US-09-745-506-74 (1-1553) x A69954 (1-373)

Qy 260 CTCCTTTCCTCTGGAATGACTTTCATCCCTCTGCTGCTGAGACTTGGACAAATGTT 319

Db 10 IleIleGlnIeuPheGluGluGlnIeuPheSerProIySalIyAlaValGluGlyAspIySile 29

Qy 320 GGATTA-----CTGCTGGAACCAAGCCACACATCACTGTAATACACTTTC 367

||||||| ||| ::|||

Db 30 GlyLeuGlnIleGlyThrLeuAsnLysPro-----IleLysAsnValMet 44

Qy 368 STGASCAATGATGCTGAGAGAGTATGTAGAGAGGTGCTGCAAAAGAGGCAAGCTTC 427

Db 45 ValThrLeuAspValIleGlnSerValIleAspIleValIleGlyLysGluValAspLeu 64

Qy 428 ATTCTCTCTCAACATCCGCTATCTCTCCGACCCATGAAAGCGCATTAACCTGGACACATGG 487

Db 65 IleIleIleAsnHisProProlIlePheArgSerLeuLysHisIleSerThrAspGlnPro 84

Qy 488 AAGAGAGCGCCGTGGATCCGGCTGTGGAGAAAGAGAGCGGTATCTATCTGCTCCATACA 547

Db 85 AlaLeuArgLeuIleGlyLysCysLeuLysHisAspIleAlaValIleAlaIleAsnHisThr 104

Qy 548 GCSATGATGCTGCGCCCGCCAGGCGCTCAACAACATGATGGCTTAAGAGGCTTGAGCTTGT 607

Db 105 AsnLeuAspValAlaAspGlyLysValAsnAspLeuIleAsnIleAlaIleGluIleSer 124

Qy 608 ACSTCCAGGCCCATATACATCTTCCAAAGCTCCCAACTAC----- 646

Db 125 GluThrGluValLeu-----AlaProThrThrAspProLeuLysLysLeu 140

Qy 647 -----CCTACAGAG-----GGAAAC----- 661

Db 141 AlaValIleValProLysGluIleValGluGluGlnValAlaIleAlaIleGluAlaIleGly 160

Qy 662 -----CACCGAGTAGAATTCACGTTAACTACACCCCAAGCACTGCAAAAGTCATGCT 715

Db 161 AlaGlnHisIleGlyGluIlePro-----SerHisGlyAlaPheSerSerGluGlyIleGly 178

Qy 716 GCACTGAAGAGGAATTCAGCGGTCTTCTGTCACATCTTTTCTGCTAGACACTGTAATAGAG 775

Db 179 SerPheLysProLeuAspGlyAlaLys-----ProPheIleGlyGluValGlyLeu 196

Qy 776 GAA-----CAACACCGATTAACTGAAATGTACTGCAAGCGCTTGATGCAAGTGC 826

Db 197 GluLeuValAlaHisGluValAlaArgLeuGluThrValPheProLysSerValGluLysAlaVal 216

Qy 827 GTAGATTTTCTTCTCCGCAACAACACTTATATCAGAAAGAG-----GAAATCTGTCA 880

Db 217 IleAsnAlaMetIleLysSerHisPro-----TyrIleGluValAlaIleTyrAspIleTyrPro 235

Qy 881 CTGGAGAAAGCTTTGCTTATACATACGTGAATGGAGTGGAGCGTTATGCAACACTGATGAATCT 940

Db 236 ValGlnGln--ThrProAlaGluLysGluValGluIleTyrValGlyThrLeuLysAsnGlu 254

Qy 941 GTCCTCCCTGGACACCATGATGATGCAATAAAAGACACCTAAACATATGCTCATATGCC 1000

Db 255 MethTrpLeuLysGluPheAlaLeuPheAlaLysAspLysLeuAspValAsnGlyValArg 274

Qy 1001 TTAGACCCCTGGGCGTGGAGAGAACTTAGAGTCTCAAGTCAAAAGTCTGGCCCTGTGCT 1060

Db 275 Met-----ValGly--AspAlaAspSerMetValLysLysValAlaValLeuGly 290

Qy 1061 GGTCTCTGGAGACACGCTTCTGCAAGGCTGTAG-----GCTGACCTTTTACCTACA 1111

Db 291 GluAspGluAsnLysTyrIleHisIleAlaLysArgLysGluAlaAspValIleTyrValThr 310

Qy 1112 GGTGAGATGTCCCATGATACCTTTGGATGCTGCTCCCAAGAAATTAATGTCAATGCTC 1171

Db 311 GluAspLeuTyrPheHisValAlaHisAspAlaMetLeuGlnLysLeuAsnValValAsp 330

Qy 1172 TGTAACACACACACACATGAGAGAGGCTTCTTCTGACCTTCAGATATGCTGATCT 1231

Db 331 ProGlnHis--TyrAlaGluLysIleMetLysLysIleValAlaIleThrArgLysLeuThrSer 349

Qy 1232 CACTTGGAGAAATAAG-----ATAATATTAATCTATACAGACTGACAGGAGCCT 1282

Db 350 MetCysAsnAspLysLysPheGlyValAlaAsnIlePheValSerGluThrAspThrAsnPro 369

RESULT 4  
889936  
Conserved hypothetical protein SA1388 [Imported] - Staphylococcus aureus (strain N315

[illegible]

Db 201 MetLeAsprLaTyGrGlnLys-----SerArgSlaGlu 211  
 : : : : : ||||| : : : : :  
 QY 851 CAATTTNATTCAGAA-----ACGGAAATTCTGTCACTCGAG 886  
 ||||| : : : : :  
 Db 212 GlnLeuIleLysGlnTyHnIsProTylGlnThrProValPheAspRheIleGluIleLys 231  
 : : : : :  
 QY 887 AAGCCTTGCTCTCATACATCTGAATGGGAGGGTATGCACACACATGTAATGTGTCTCC 946  
 : : : : : ||||| : : : : :  
 Db 232 GlnThrSerLeuTyG----GlyLeuGlyAlaMetLalGluValAlaPasnGlmMetThr 249  
 : : : : : ||||| : : : : :  
 QY 947 CTGGCAACCATGATGATGATGATAAAGAACCCTAAACTATCTCATATGTGGCTTAGGCC 1006  
 ||||| : : : : :  
 Db 250 LeuGlnAspRheAlaAlaAspRleLysSerLysLysLysLysLysProSerValArgPhe--- 268  
 : : : : : ||||| : : : : :  
 QY 1007 CTTGGGGTGGGGAGAACCTTGAAGTCTCAAGTCAAAGTCTGGGCCCTGTGCTGGTTCT 1066  
 ||||| : : : : :  
 Db 269 -----ValGlglyLysr---AsnGlnLysIleLysArgIleAlaIleIleGlglySer 285  
 : : : : :  
 QY 1067 GGG-----AGCAGCGTCTGGAGGGGTGTGAGCGTGAGCTTACTTACCTACA 1111  
 ||||| : : : : :  
 Db 286 GlyIleGlyTyGlnTyGlnAlaValGlnGlnGlnLy-----AlaAspValPheValThr 303  
 : : : : :  
 QY 1112 GGATGAGATGTCCTCCATGATATATCTTGGATGCTGCTCTCCAGGAATAATATGTCATCTC 1171  
 ||||| : : : : :  
 Db 304 GlyAspRleLysShHisAlaSprAlaLeuAspRlalySlenHisGlyValAlaSnleuIleAsp 323  
 : : : : : ||||| : : : : :  
 QY 1172 TGTGACACACAGCACATGTAACGAGGCTTTCTTCTGACCTTCAGATATGCTGGATTCT 1231  
 ||||| : : : : :  
 Db 324 IleAsnHis---TyrSerGlnTyGValMetLysGlnGlyLeuLysThrLeuLeuMetAsn 342  
 : : : : : ||||| : : : : :  
 QY 1232 CACTTGGAG---AATPAGATTAATATTTATCTCTA-----TCAGAGACTGACAGGAGCCCT 1282  
 ||||| : : : : :  
 Db 343 TrpPheAsnIleGlnLysIleAsnIleAspValGlnAlaSerThrIleAsnThrAspPro 362  
 : : : : :  
 QY 1283 CTTTCAGTGTGTA 1294  
 ||||| : : : : :  
 Db 363 PheGlnTyRile 366

RESULT 5  
G97060  
uncharacterized protein of YbgJ/Acr family [Imported] - Clostridium acetobutylicum  
C:Species: Clostridium acetobutylicum  
C:Date: 14-Sep-2001 #sequence\_revision 14-Sep-2001 #text\_change 14-Sep-2001  
C:Accession: G97060  
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; L.  
. J. Bacteriol. 183, 4833-4838, 2001  
J.: Daly, M.D.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.  
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium  
A:Reference number: A96900; MUID:21359325; PMID:21359325  
A:Accession: G97060  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-268 <R>  
A:Cross-references: GB:AE001437; PIDN:NAK9274.1; PID:G15024233; GSPDB:GN00168  
A:Experimental source: Clostridium acetobutylicum ATCC624  
C:Genetics:  
A:Gene: CAC1303

Alignment Scores:

Pred. No.:	3_97e-14	Length:	268
Score:	264.00	Matches:	75
Percent Similarity:	39.20%	Conservative:	63
Best Local Similarity:	21.31%	Mismatches:	118
Query Match:	9.47%	Indels:	96
DB:	2	Gaps:	6

US-09-745-506-74 (1-1553) x G97060 (1-268)

QY 245 ATGATTTGAAGAGGCTCTCTCTTCCTTCCTGATGACATTGGATCCCTCTTGCTGAG 304  
 : : : : : ||||| : : : : :  
 Db 3 LeuLysValLysAspLeucyAsnIleIleGlnAspRheAlaProIleSerLeuLysGlu 22

OY 305 GGTGGGACATGTTGGATTACTGCTGGGAACCAAGCCACCACACTACTGTAATTAACCTC 364  
 Db 23 AspheAspsnValIGlyLeuMetValIGlyAspArgGluLa---SerValaPaIaIe 41  
 OY 365 TTCCCTGCACATGACCTGCAGTAGAGAGATGATGAGAGGGTGTGCAAAAGAGGACAG 424  
 Db 42 MethrAlaLeuAspCysThrMetAspValIleAspGluLaIleGluLysAnCysAsn 61  
 OY 425 CTCATTCTCTCTACCACTCCGCCCTATTTCTCCGACCCATGAAAGCCATTAACCTGGACACA 484  
 Db 62 MetIleIleThrHisHisProIleLeuPheLysLysProSerLysIleThrMetAspThr 81  
 OY 485 TGAAGAGGACCCCTGGGATCCGGCTCTGGAGACAGAGTCGGTATCTACTCTCCCTCAT 544  
 Db 82 LeuLeuGlyLysIleIleLysIleIleSerHisnIleAsnValIYrSerAlaHis 101  
 OY 545 ACAGCCATGATGCTGGCGCCCGGCGTCACACAACTGTTGGCTAAAGGCTTGGAGCT 604  
 Db 102 ThrAsnLeuAspSerValLysAspGlyIleAsnAspAlaValValAsnIleLeuGly--- 120  
 OY 605 TGTACCTCCAGGCCCCATACATCCTTCCAAAGCTCCCACTACCTTCACAGAGGAACAC 664  
 Db 120 ----- 120  
 OY 665 CGAGTAAATTCACGTTAACTACACCCAGACACTGGACAAAGTCATGCTGCAGTGA 724  
 Db 120 ----- 120  
 OY 725 GGAATTACGGGTGTTCTGTCACTTCTTTTCTGCTAGAGACTGTAATGAGAAACA 784  
 Db 120 ----- 120  
 OY 785 CGGATTATCTGAATTGTACTCAGAAAGCTTTGATGAGGTGGTAAGATTCTTCCCGG 844  
 Db 120 ----- 120  
 OY 845 AACAAACAACTTTATTCAGAAAGCGAAATTCCTGCTACGTGAGAAAGCCTTTGCTTACAT 904  
 Db 121 -----PheAspLysSerSerIleLeuAlaLysAsnLysAlaValLysGlu 136  
 OY 905 ACTGAATGGAGCGTTATGCACACCTGGAGGAATCTGTCCTCCGCGCAACCATGATTGAT 964  
 Db 137 AlaeIlyIleGlyArgValValIGluLeuGluGlnIsmMetThrLeuLysGluLeuCysAsp 156  
 OY 965 CGAATTAAGAGACACCTAAACATATCTCATATTCGCTTACCCCTTGGGGTGGGAGAAC 1024  
 Db 157 ArgValLysGlnSerPheLysIleGlnSerLeuArgLysCysGlyAspGluAspLysLys 176  
 OY 1025 TTAGAGTCTCAAGTCAAGTCGTCGCGCCCTGTGCTGCTGCTGGAGACAGCTTTCGAC 1084  
 Db 177 IleHisSer-----PheAlaValIleAsnLysSerLysGlnAspPheGlu 192  
 OY 1085 GGTGTTAG-----GCTGACCTTTTACTCTCAGAGTCGATGTCGATTCATCT 1135  
 Db 193 GluAlaArgLysArgLysAlaAspCysIleIleThrGlyAspHisSerLysIYrHisVal 212  
 OY 1136 TTGGATGCTCTCTCCCAAGAAATTAATATGTCATCTCTGTGACACACAGCAACACTGACGA 1195  
 Db 213 SerAspLysArgnGluMetAsnIleAlaValIleAspAlaGlyHisPheGlyThrGluTrp 232  
 OY 1196 GCGTTTCTTCTGACCTTCGAGATATGCTCGATTCTTCACCTTGGAGAAATAA3---ATRAAT 1252  
 Db 233 ProSerValAlaValMetSerLysLysLeuGluGlyAlaLeuHisLysMetGlyIleAsn 252  
 OY 1253 -----ATTATCTTATCAGAGACTGACGAGGAGCCCT 1282  
 Db 253 ThrProIleLeuValSerGlnAsnAsnIleAspPro 264

[illegible]

```
Db 261 LysGlnValLysLysAlaPheAspValProThrAlaArgVal-----ValGly--- 276
QY 1022 ACCTTAAGTCTCAAGTCAAAAGTGGCCCTGTGTGCTGTCTGGAGCAGCGTTTCG 1081
Db 277 SerLeuGlnThrGlnIleArgLysValAlaValLeuGlnLysAspGlyAsnLysTyrMet 296
QY 1082 CAGGGTGT-----GAGCGTGACCTTTACTCTACAGAGTGAGATGCTCCATCATAT 1132
Db 297 AlaHisAlaLeuArgLysGlyAlaAspValIleValThrGlyAspValTyrHisVal 316
QY 1133 ACTTGGATGCTGCTCCAGGAATATGATCATCTCTGTGAACAGACAGACCTGAA 1192
Db 317 ProHisAspAlaLeuMetAspGlyLeuAsnIleValAspProGlnHis---AsnValGlu 335
QY 1193 CGAGCGTTTCTTTCTGACCTTCGAGATATGCTGATTCCTCACTTGGAGAAATAG----- 1246
Db 336 LysIleMetLysGlnGlyValLysGlnLysLeuGlnLysLeuLysAspLysLysTyr 355
QY 1247 ---ATTAATATATCTATCTACAGACATGACAGGACCTTTACAGGTGTA 1294
Db 356 AspThrGlnValAlaIleAspSerValHisThrAspProPheThrPheIle 372

RESULT 7
AH1618
conserved hypothetical protein lin1489 [imported] - Listeria innocua (strain Clp11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AH1618
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker,
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.;
Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kretz, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehlund,
A.; Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AH1618
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-373 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC96720.1; PID:gl6413962; GSPDB:GN00178
C:Experimental source: strain Clp11262
C:Genetics:
A:Gene: lin1489

Alignment Scores:
Pred. No.: 2,166-13 Length: 373
Score: 256.00 Matches: 86
Percent Similarity: 42.93% Conservative: 78
Best Local Similarity: 22.51% Mismatches: 152
Query Match: 9.18% Indels: 66
DB: 2 Gaps: 11

US-09-745-506-74 (1-1553) x AH1618 (1-373)
QY 272 TTGAATGACTTGCATCCCTCTGTTGCTGAGACTGGACATGTTGGATTTACTGTG 331
Db 13 MetGlnLysIleAlaProLysLysLeuAlaMetGlnLysAspProIleGlyLeuGlnVal 32
QY 332 GAACCAAGCCGACCAACATCTGTAATACACTCTCTGACCAAGACCTGACATGAGAA 391
Db 33 GlysAspLeuSerLysLysValAlaGlyLysValMetPhe---ThrLeuAspValIleuGlu 51
QY 392 GTGATGAGAGAGTGTCTCAAAAGAGAGAGACCCATTTCTCTCTACATCCGCTATC 451
Db 52 ValValAspArgLysAlaIleGlnLysArgValAspLeuIleIleAlaHisHisProHeu 71
QY 452 TTCGACCCATGAGACGCAATACCTGGAACACATGGAAGAGAGCGCTGGTATCCGGCT 511
Db 72 TyrArgProThrGlnHisIleAspThrThrThrLysGlnGlyLysMetIleLysLysLeu 91
QY 512 CTGGAACACAGAGTGTATCTACTCTCTCATACAGCTATGATGCTGCGCCCAAGGC 571
```

```
Db 92 IleLysHisAspIleThrValAlaAlaHisHisThrAsnLeuAspIleAlaGlnGly 111
QY 572 GTCAACAACCTGGTTGGCTAAAGGCTTGGACCTGTACTCTCCAGGCCATACATCTTCC 631
Db 112 ValAsnAspIleLeuAlaAspLeuLeuHisIleuGlnAspThrThrMetIleGluThr 131
QY 632 -----AAAGCTCCCACTACCTACAGG----- 655
Db 132 TyrThrGlnProTyrCysLysIleAlaValIleAlaProGlnAsnGlnLeuGlnSerVal 151
QY 656 -----GGAACCCAGAGTAAATGATCAAGCTTAATCACTACACCCAGACCTG 700
Db 152 ArgLeuAlaLeuValAsnAsnGlyAlaGlyGlnIleGlyThrAsnTyrThrGlu----- 169
QY 701 GACAAAGTATGTCTGACGATGAAGAGAAATGACAGGTGTCTGTCTCTTTTCTGCT 760
Db 170 -----CysThrPheHisIleThrGlyLysLeuSerPheLysProGlyThrAspAlaAsn 187
QY 761 AGGACTGGTAATGAGAACCAACAGCGATTAATCTG----- 796
Db 188 ProThrIleGlyGlnLysGlyThrLeuThrSerIleThrProGlnValLysIleGlnAlaIle 207
QY 797 -----AATTGACTCAGAGGCTTGTATG----- 820
Db 208 PheProGlnTyrLeuThrGlnThrIleThrLysAlaValLysIleAlaHisProTyrGlu 227
QY 821 ---CAGGTGATAGATTTTCTTCCCGAACAACATTTATACAGACGGAATCTGT 877
Db 228 GluProAlaIleAspValTyrThrLeuGlnThrGlnThrTyrLysGlu----- 243
QY 878 TCACGTGAGAAACCTTCTCTCATACCTGGAAGGAGCGTTATGACACCTGATGGA 937
Db 244 -----GlyLeuGlnArgValGlyThrLeuProLys 253
QY 938 TCTGTCTCCGCAACCATGATGATGATGAATAAAGACACCTAAATCTTCATATT 997
Db 254 LysIleSerMetValSerPheIleAspLysLeuLysThrAlaPheAlaIleAspAsnVal 273
QY 998 CGCTTAGCCCTTGGGCTGGGAGAACCTTACAGTCTCAAGTCAAGTCAAGTCTGGCCCTGT 1057
Db 274 ArgPhe-----ValGly---AspLeuLysAlaAsnValGlnValAlaIleIle 289
QY 1058 GCTGCTTCTGGAGCAGCGCTTCTG-----CAGGCTGTGAGCGTACCTTTACCTC 1108
Db 290 GlyGlnAspGlyAsnLysPheIleHisGlnAlaLysAlaThrGlyAlaAspValPheIle 309
QY 1109 ACAGTGAAGATGCCATCATGATCTTGTGATGCTGCTCCAGGAATAAATGTCATC 1168
Db 310 ThrGlyAspValTyrTyrHisThrAlaHisAspLeuAlaIleAsnLeuProThrIle 329
QY 1169 CTCTGTGAACACAGC-----AACACTGAACGAGGCTTCTTCTGACCTTCGAGATATG 1222
Db 330 AspAlaGlnHisAsnIleGlnLysValMetLysGlyTyrLeuLysAsnLysMetLysGlu 349
QY 1223 CTGGATTTCTACTTGGAGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1282
Db 350 GlnAlaLysIleLeuAspLysGlnAlaGlnPheIleValSerGlnValAsnThrAspPro 369
QY 1283 CTTCAG 1288
Db 370 PheGln 371

RESULT 8
AD1256
conserved hypothetical protein lmo1452 [imported] - Listeria monocytogenes (strain EG
C:Species: Listeria monocytogenes
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloec
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl,
Science 294, 849-852, 2001
```



```

Db      78  TyrProAlgaIaPheIleLysGluMetIleTyrLysAsnIleSerLeuIleSerMetHis 97
OY      545  ACAGCCATGATGCTGCGCCCGAGGGCGTCACAAACACTGGTGGTAAAGGGCTTGAGCT 604
Db      98  ThrAsnTyrAspLeuSer----- 103
OY      605  TGTACCTCCAGGCCATACATCCTTCACAAAGCTCCCAACTACCTACAGAGGAAACGAC 664
Db      103  ----- 103
OY      665  CGAGTGAATTCACGTTAATCTACACCCAGACGCTGCACAAAGTATGTCTGCAGTGA 724
Db      104  -----HisLeuAsnThrTyrPheThrGluGlu----- 112
OY      725  GGAATTGACGGTGTCTGTCACTTCTTCTGCTAGACACTGTATAGAGGAACAACA 784
Db      113  -----IleLeuGlyPheLys----- 117
OY      785  CGGATTAACTGGAATTGACTCAGAAAGCCTTGATGACAGGTGTAGATTCTTTCCTCCGG 844
Db      117  ----- 117
OY      845  AACAAACAACCTTATCAGAACGGAATTCGTCTGCTGAGAACCTTTGCTTCTACAT 904
Db      118  -----IleSerPheLysAspLysPheLeuIleTyr 127
OY      905  ACTGGAATGGACGGTTATGACACTGTGATGATCTGCTCCCGCAACCATGATTGTAT 964
Db      128  -----ValGluAsnSerMetSerPheGluAlaLeuGlyAsp 139
OY      965  CGAATAAAGACACCTAAACATATCTCATATTCGCTTAGCCCTTGGGGTGGGGAAC 1024
Db      140  TrpValLysLysLeuLeuAsnLeuGlnIleLeuArg-----ValSerAspCys 155
OY      1025  TTAGAGTCTCAAGTAAAGTGTGGCCCTGTGTGCTGTGCTGAGAGCAGCTTCTGCAG 1084
Db      156  GlyLysLysAspLysPheLysArgIleAlaIleCysThrGlySerGlyLysPheLeuSer 175
OY      1085  GGTGTGAGGCTGACCTTTTACCTCCACAGGTGATGATGCTCCATGATGATCTTGGATGCT 1144
Db      176  LysValAspAlaAspCysPheLeuSerGlyAspPheLysTyrHisGlnAlaLeuGlnAla 195
OY      1145  GCTTCCCAAGCAATAAATGCTATCTCTGTGACACAGACACAGTGAAGAGGCTTT--- 1201
Db      196  LeuSerAsnGlnIleSerLeuIleAspLeuGlnHisPheGluSerGluArgTyrPheSer 215
OY      1202  -----CTTTGTACCTTGCAGATATGCTGATTCCTCACTTGAGAAATAAGATTAAT 1252
Db      216  GlnCysLeuAlaLysAspLeuLysAsnLeu----- 225
OY      1253  ATTATCTATTCAGAGACTGACAGGAGCCCTTTCAGTGCTA 1294
Db      226  -----ProLeuGlnValIle 230
RESULT 10
B70777
Hypothetical protein RV2230c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70777
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; M01D:98293587; PMID:9634230
A:Accession: B70777
A:Status: Preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-379 <CDL>
A:Cross-references: GB:I270692; GB:AL123456; NID:93261567; PIDN:CAA94653.1; PID:e235171.

```

```

A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2230c

Alignment Scores:
Pred. No.: 4,966-10 Length: 379
Score: 216.50 Matches: 95
Percent Similarity: 44.01% Conservative: 63
Best Local Similarity: 26.46% Mismatches: 158
Query Match: 7.76% Indels: 43
De: 2 Gaps: 15

US-09-745-506-74 (1-1553) x B70777 (1-379)
OY      299  GCTGAGAGTTGGAGACATGTTGATTACTGTTGAGGACCAAGGCCACACATCTGTAAT 358
Db      21  AlaGlnSerThrPaspSerValGly---LeuValCysGlyAspProAspValValasp 39
OY      359  ACACCTTCCCTGACCAATGACCTGACTGAGAGATGATGAGAGAGCTGCTGCAAAAGAG 418
Db      40  SerValThrValAlaValAspAlaThrProAlaValAlaValAspGlnValProGln----- 57
OY      419  GCAGACGTCATCTCTCTGCTACCATCCGCTATCTCCGACCCATGAGGACGATACCTGG 478
Db      58  AlaGlyLeuLeuLeuValHisHisProLeuLeuLeuAlaGlyValAspThrValAlaAla 77
OY      479  AACACATGGAAGAGACCGCTGATCCGGGCTGTGAGAACAGAGCTGATCTACTCT 538
Db      78  AsnThrProLysGlyValLeuValHisArgLeuIleArgThrGlyArgSerLeuPheThr 97
OY      539  CCTCATACAGCCTATGATGCTGCGCCCGCCAGGCGCTCAACACTGTTGCTTAAGGCTT 598
Db      98  AlaHisThrAsnAlaAspSerAlaSerProGlyValSerAspAlaLeuAlaHisAlaVal 117
OY      599  GGAGCTTGACCTCCAGGCGCCATACATCTCCAAAGCTCCCAAC----- 643
Db      118  GlyLeuThrValAspAlaValAlaLeuAspPro-----ValProGlyAlaAlaAspLeuAsp 135
OY      644  -----TACCTACAGAGGGAACCAACCGAGTA-----GAA 673
Db      136  LysThrValIleTyrValProArgGlnAsnSerGlnAlaValAlaValAlaValPheGln 155
OY      674  TTCAAACGTTAACTAACCCCAAGACCTGGACAAAGTCATGCTGCAGTGAAGGATTT--- 730
Db      156  AlaGlyAlaGlyHisIleGlyAspTyrSerHisCysSerTrpSerValAlaIleThrGly 175
OY      731  -----GACGGTCTTCTGTGCACTTCTTTTCTGCTAGACACTGTAATGAG 775
Db      176  GlnPheLeuAlaHisAspLysAlaSerProAlaIleGlySerValGlyThrValGlnArg 195
OY      776  ---GAACAAACAGGATTAATCTGAATTGTACTCAGAAAGGCTTTGATGCAGGTGCTAGAT 832
Db      196  ValAlaGlnAspArgValGlnValAlaAlaProAlaArgAlaArgAlaGlnValAlaLeuAla 215
OY      833  TTTCTTCCCGGAGCAACAACTTAATCAGAGAG-----GAAATCTGTCACTGGAG 886
Db      216  AlaMet---ArgAlaAlaHisProTyrGluGluProAlaPheAspIlePheAlaLeuVal 234
OY      887  AAGCTTTGCTCTTACATCTGATGAGGAGCGGTTATGACACACTGATGATCTCTCC 946
Db      235  ProPro-----ProValGlySerGlyLeuGlyArgIleGlyArgLeuProLysProGluPro 253
OY      947  CTGGCAACCATGATGATGCAATAAACACACCTAAACCTAAATCTATTCATATTCGTTAGCC 1006
Db      254  LeuArgThrPheValAlaArgLeuGlnAlaAlaAlaLeuProProThr-----Ala 269
OY      1007  CTGGGGTGGGGAAGAACCTTAAGTCTCAAGTCAAGTC-----GTGGCCCTGTGTGCT 1060
Db      270  ThrGlyValAlaArgAlaAlaGlyAspProAspLeuValSerArgValAlaValCysGly 289
OY      1061  GGTTCGGGAGGACGGTTCAGAGGTGTGAGGCTT-----GACCTTAACTGCACA 1111
Db      290  GlyAlaGlyAspSerLeuLeuAlaThrValAlaAlaAlaAspValGlnAlaTyrValThr 309

```



OY 1112 GGTGACATGGCCCATC-----GATACCTTTGGATGCTGCTGCCAAGAAATAAATGTC 1165  
          :::      |||||          ||          |||||          :::      :::  
Db 310 AAlaspleuraighishisprolialaspelunhisysargalasergin--valAlaleu 328  
OY 1166 ATCCCTGTGAACAACAACACTGAACAGCGCTTCTTCTTGACCTTCGASATATGCTG 1225  
          ||          ||          :::||||          :::          :::      |||  
Db 329 Ilaspvayalainstrialaserglupheprotyrpsylglnalaiaalsjvalleu 348  
OY 1226 GATTTCATCTGGAGATTAAGATAATATATATCTATACAGAGACTGACAGSGACCT 1282  
          |||||          :::      :::      :::      :::      |||  
Db 349 ArgSerIshfheglyalaserleuprovalaryValcysthrIllecysthraspro 367

RESULT 11  
AI2207  
hypothetical protein alr3216 [imported] - Nostoc sp. (strain PCC 7120)  
C/Species: Nostoc sp. PCC 7120  
A/Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120  
C/Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 09-Dec-2002  
C/accession: AI2207  
R/Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kurlitz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi,  
Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8: 205-213, 2001  
A/title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anaer  
A/reference number: AB1807; MUID:21595285; PMID:11759840  
A/accession: AI2207  
A/status: preliminary  
A/molecule type: DNA  
A/residues: 1-263 <KDR>  
A/cross-references: GB:BA000019; PIDN:BAB74915.1; PID:g17132311; GSPDB:GN00179  
A/experimental source: strain PCC 7120  
C/genetics:  
C/gene: alr3216

Alignment Scores:	
Pred. No.:	1,7e-09
Score:	209.50
Percent Similarity:	35.33%
Best Local Similarity:	22.51%
Query Match:	7.51%
DB:	2
Length:	263
Matches:	79
Conservative:	45
Mismatches:	126
Indels:	7
Gaps:	6

US-09-745-506-74 (1-1553) x AI2207 (1-263)

QY	245	ATGATTTTGAAGGCGCTCTCTCTCTCTCTGATGACTTTGGATTCCTCCCTGTTGGTGTGAG	304
Db	1	Mettysilealaaspleuilethtrpphegiusertrpalaasnprolatrrpcysglu	20
QY	305	AGTTGGGCAATGTTGGATTACTGGTGGACCAACGCCCAACATCTGTAAATACACTC	364
Db	21	Sertrpapaancysgilytrpelnileguproglytleuonileglunalaargval	40
QY	365	TTCTGACCAATGACCTGACTGAGAGAGGAGGAGGAGTGTGCAAAAAGAGCGAGAC	424
Db	41	LeuvalCysleuthrProthrleualaValmetglunalailelealeuglnalaasn	60
QY	425	CTCATTTCTCTACACTCGCGCTATCTTCCGACCATGACCGCATTAACCTTGACACAC	484
Db	61	LeuilepmealanhshisProleuilepbeserProprolySerleuulrgaTgylglu	80
QY	485	TGGAAGGACGCTGGTGGATCCGGCTGTGGAGAACAGAGTGCGGTATCTACTCTCCAT	544
Db	81	AlailealaaspmelaAlargleualaPhethrlyasnilleglyleuTserlalanhs	100
QY	545	ACAGCCTATGATGCTGCCGCCCGGCGGCGTCAACACTGGTGGCTTAAGGCGTGGAGCT	604
Db	101	ThrsanpheaaprglnValgluasplythrAlaasprvalLeualaeglInlleuuglyleu	120
QY	605	TGTACTCTCAAGGCCCATACATCTTTCCAAAGCTCCCACTACCTACAGAGCAACAC	664
Db	121	-----LysasprvalAlaProIlevalProthrInglugly-----	132
QY	665	CGAGTAGACTTCACAGCTTAACACTACACCCCAAGACTGGACCAAAAGTCACTGTCTGCAGGAAA	724

Db	132	-----	132
OY	725	GGAATTGACGGTCTTCTGTCACTCTTTTCTGTAGACGTGTAATGAGAACAAACA	784
Db	132	-----	132
OY	785	CGGATTATCTGAACTTACTCAGAAAGCTTTGATGAGGTGGTAGATTTCTTCCCG	844
Db	132	-----	132
OY	845	AACAAACAATTATCAGAGACGGAATTCGTCACTGGAGAACCTTTCCTTACAT	904
Db	132	-----	132
OY	905	ACTGGAATGGGACGGTTATGCACACTGGATGAACTCTCTCCCTGGCAACCATGATTGAT	964
Db	133	Leu11Tyrg1Valg1Valg1Leu1Leu1Asp1Pro1Phe1Leu1Ser1Gln1Glu1Leu1Thr	152
OY	965	CGAATTAAGAGACACCTAAACCTATCTCAATATTCGCTTACACCTTGGGGGGGAGAAC	1024
Db	153	Val1Leu1-----Thr1Arg1Leu1Ala1Pro1Phe1Asp1Leu1Phe1Ser1Pro1Thr1Ala1Asp	170
OY	1025	TTAGAGTCTCAAGTCAAAAGTCGTGGCCCTGTGTCTGGTCTTCTGGGAGACACGTTCTGCAG	1084
Db	171	Leu1Gln1Gln1Leu1Leu1Ser1Arg1Val1Ala1Val1Leu1Gly1Ser1Gly1Ala1Gly1Leu1Ser	190
OY	1085	GGGTCTGAG-----GGTGACCTTACTCTCAGAGTGAGATGTCCATCAATGATACT	1135
Db	191	Ala1Val1Ala1Gln1Thr1Gly1Ala1Gln1Ala1Tyr1Leu1Thr1Ser1Asp1Cys1Lys1Phe1Leu1Gln1Phe	210
OY	1136	TTGGATGCTGCTTCCCAAGAAAGTAATGTATCTCTGTATACACACACACACTGAGCA	1195
Db	211	Gln1Gln1Ser1Arg1Asp1Arg1Gly1Leu1Leu1Gln1Leu1Ser1Ala1Gln1Tyr1Ala1Thr1Gln1Arg	230
OY	1196	GGCTTTCTTTCTGACCTTCGAGATATGCTGGATTCTCAGTTGGAGAAATAAGATAAT--	1252
Db	231	Pro1Ala1Cys1Asp1Arg1Leu1Ala1Gly1Leu1Arg1Ser-----Leu1Asn1Leu	245
OY	1253	-----ATTATCATCAAGACACTGCAGAGGAC	1279
Db	246	His1Tyr1Val1Gln1Leu1Ser1Asn1Gln1Asp1Gln1Asp	256

RESULT 12

C:\Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 03-Aug-2001  
C:\Accession: D95187  
D95187/ conserved hypothetical protein SPI609 [imported] - Streptococcus pneumoniae (strain T

R. Tetteijn, H. Nelsson, K.E.; Paulsen, I.T.; Elsen, J.A.; Reed, T.D.; Peterson, S.; H  
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzappl  
nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A: Authors: Lofstis, B.J.; Yand, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris  
A: Title: Complete Genome Sequence of a Virulent Isolate of *Streptococcus*  
A: Reference number: A95000; MUID:21357209; PMID:11463916

A: Accession: D95187  
A: Status: preliminary

A: molecule type: DNA  
A: Residues: 1-265 <KUR>  
A: Cross-references: GB:A:E005672; PIDN:ANK75693.1; PID:g14973101; GSPDB:GN00164; TIGR:  
D: RefSeq accession number: NC\_007804

A: Experimental Source: Strain 11694  
C: Genetics:  
A: Gene: SPI609

Alignment Scores:

Pred. No.:	3.53e-09	Length:	265
Score:	203.50	Matches:	67
Score:	27.10	Compositions:	53

Percent Similarity:	37.10%	Conservative:	02
Best Local Similarity:	19.31%	Mismatches:	121
Query Match:	7 30%	Indels:	97

DB:	2	Gaps:	7
-----	---	-------	---

US-09-745-506-74 (1-1553) x D95187 (1-265)

QY 260 CTCCTTCTTCCTGTAATGACTTTCATCCCTCGTTGCTGAGAGTTGGACAATGTT 319  
Db 6 Valileglnalatyrglnalaphecysproglnglnuphesermetgluglyaspseryc 25

QY 320 GGATTACTGTGGAGACCAAGCCACCATCTGTAATATACCTCTTCCTGACCAATGAC 379  
Db 26 Glyleuglnile--GlyThrleuasplysglylelglnargvalmetvalaleuasp 44

QY 380 CTGCTGAGAGAGGAGGAGGTGCTGCAAAAAGAGGACCAATCTCTCCAC 439  
Db 45 Ilearglgluglnthrvalalaglnalaleglulysglvalaspleuilelevallys 64

QY 440 CATCCGCTATCTTCGACCCATGAGCGCATTAACCTGGAACACATGGAAGAGCGCTG 499  
Db 65 Hislaproilepheargproilelysaspleuulaserargproglasnglnle 84

QY 500 GTGATCCGGGCTGTGGAGACAGAGTCGTACTCTCTCATACAGCTTATGATGCT 559  
Db 85 Tyrilleaspleuilelyshlaspsilealavaltyrvalserhsthrasnilleasple 104

QY 560 GCGCCCGAGGCGTCACAACTGTTGGCTAAGGGCTTGAGCTTGACTCCAGGCC 619  
Db 105 Valgluasnnglyleuasnaspttrphecysglnmetleu 117

QY 620 ATACATCTTCGAAAGCTCCCACTACCTACAGAGGGAACACCGAGTAATTCAC 679  
Db 117 ----- 117

QY 680 GTTAACACACCCAGAACCTGAGAACATGTCGTCTGAGTGAAGAAATGAGCGTGT 739  
Db 118 -----Glylle 119

QY 740 TCTGTCACTTCCTTTTCTGTAGAGACGTGTAATGAGAACAAACAGGATTATCTGAT 799  
Db 120 Gluglnthrthrtyrleuglnlthrclgylprogluarg 132

QY 800 TGTACTCAGAAAGCTTTGATGACAGGTGTAGATTTCTTCCCGAACAAACATTTAT 859  
Db 132 ----- 132

QY 860 CAGAAAGCGAAATTTCTGTACATGAGAGGCTTTGCTTACATCTGGAAGGAGCG 919  
Db 133 -----Glyllelytrg 136

QY 920 TTATGACACTGTGATGATCTCTCCCTGCAACCATGATGATGATGATGATGATGAT 979  
Db 137 Ilegllyasnile--GlnProglInthrphetrpgleuualaglnlvallyselval 155

QY 980 CTAAAGCTATCTATTCCTGAGCCCTTGAGGGGGGAGACACTTGAAGTCAAGTC 1039  
Db 156 PheaspleuaspsersleuargmetvalhstfyrnglnluaspaspleuGlnlysprole 175

QY 1040 AAAGTCGAGCGCTGTGTGCTGTTCTGAGAGACGCTTGACAGGGGTGTT----- 1090  
Db 176 Serargvalalalalecysglyserglylnserphetrtyrlyaspsalaleuualys 195

QY 1091 GAGCTGACCTTACCTCAGAGGTGAGATGTCATCATGATCTGTTGGAGCTGCTTCC 1150  
Db 196 Glylalaaspsalvaltyrlethrclgylasprlelytrhsthralaglnaspsmetleuser 215

QY 1151 CAAAGATTAATGTCATCTCTGTGACACAGCAACACTGAACGAGCTTCTTCTGAC 1210  
Db 216 Aspleuileuualaleuaspsproglyhis--Tyrillegluvalillephevalglulys 234

QY 1211 CTGAGAGTATGCTGATCTCCTGATGAGTAAG-----ATAAATATATCTTA 1261  
Db 235 IlelalaaleuuserGlntrpIysgluaspIysgltyrpsrleleaspleuupro 254

QY 1262 TCAGAGACTGACAGGAGCCCT 1282  
Db 255 Serclnalaaserthrasnpro 261

RESULT 13  
E98054  
conserved hypothetical protein sprl462 [Imported] - Streptococcus pneumoniae (strain  
C:Species: Streptococcus pneumoniae  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 22-Oct-2001  
C:Accession: E98054  
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.  
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.  
Y.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001  
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.  
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234  
A:Accession: E98054  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1265 <KUP>  
A:Cross-references: GB:AE007317; PIDN:NAL00266.1; PID:915459119; GSPDB:GN00174  
A:Genetics:  
A:Gene: sprl462

Alignment Scores:  
Pred. No.: 5.53e-09 Length: 265  
Score: 203.50 Matches: 66  
Percent Similarity: 36.68 Conservative: 62  
Best local Similarity: 18.91 Mismatches: 124  
Query Match: 7.30% Indels: 97  
DB: 2 Gaps: 7

US-09-745-506-74 (1-1553) x E98054 (1-265)

QY 260 CTCCTTCTTCCTGTAATGACTTTCATCCCTCGTTGCTGAGAGTTGGACAATGTT 319  
Db 6 Valileglnalatyrglnalaphecysproglnglnuphesermetgluglyaspseryc 25

QY 320 GGATTACTGTGGAGACCAAGCCACCATCTGTAATATACCTCTTCCTGACCAATGAC 379  
Db 26 Glyleuglnile--GlyThrleuasplysglylelglnargvalmetvalaleuasp 44

QY 380 CTGCTGAGAGAGGAGGAGGTGCTGCAAAAAGAGGACCAATCTCTCCAC 439  
Db 45 Ilearglgluglnthrvalalaglnalaleglulysglvalaspleuilelevallys 64

QY 440 CATCCGCTATCTTCGACCCATGAGCGCATTAACCTGGAACACATGGAAGAGCGCTG 499  
Db 65 Hislaproilepheargproilelysaspleuulaserargproglasnglnle 84

QY 500 GTGATCCGGGCTGTGGAGACAGAGTCGTACTCTCTCATACAGCTTATGATGCT 559  
Db 85 Tyrilleaspleuilelyshlaspsilealavaltyrvalserhsthrasnilleasple 111

QY 560 GCGCCCGAGGCGTCACAACTGTTGGCTAAGGGCTTGAGCTTGACTCCAGGCC 619  
Db 105 Valgluasnnglyleuasnaspttrphecysglnmetleu 117

QY 620 ATACATCTTCGAAAGCTCCCACTACCTACAGAGGGAACACCGAGTAATTCAC 679  
Db 117 ----- 117

QY 680 GTTAACACACCCAGAACCTGAGAACATGTCGTCTGAGTGAAGAAATGAGCGTGT 739  
Db 118 -----Glylle 119

QY 740 TCTGTCACTTCCTTTTCTGTAGAGACGTGTAATGAGAACAAACAGGATTATCTGAT 799  
Db 120 Gluglnthrthrtyrleuglnlthrclgylprogluarg 132

QY 800 TGTACTCAGAAAGCTTTGATGACAGGTGTAGATTTCTTCCCGAACAAACATTTAT 859  
Db 132 ----- 132

QY 860 CAGAAAGCGAAATTTCTGTACATGAGAGGCTTTGCTTACATCTGGAAGGAGCG 919



A: Experimental source: serovar 3; biovar 1  
 C: Genetics:  
 A: Gene: U0351  
 A: Genetic code: SGC3

## Alignment Scores:

Prog. No.:	1.62e-08	Length:	262
Score:	198.00	Matches:	77
Percent Similarity:	38.64%	Conservative:	59
Best Local Similarity:	21.88%	Mismatches:	104
Query Match:	7.10%	Indels:	112
DB:	2	Gaps:	12

US-09-745-506-74 (1-1553) x G82901 (1-262)

```

QY 245 ATGGATTGTAAGCT-----CTCCCTTCTCTGTAATGACTTTGCATCCCTCTCTGTT 298
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 4 MetaspilielysAlaGlnAspIleuAsnPhelThrLysTyrAspLeuSerLys 23
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 299 GCTGAGAGTTGGACAAATGTTGATTTACTGTGTGAGAACCAAGCCACACATCTGTAAT 358
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 24 AlaGlnAlaTyrAspLysAsnGlyLeuPheAspGlnGln-----GlnThrIleAsn 41
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 359 AACTCTCTCCGACCAATGACTGTGAGAGAGTGAAGAGAGTGTGCAAAAGAG 418
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 42 AsnValGlnIleAlaLeuAsnProIleThrAspAspValAlaAsnAspAlaIleLeuAsn 61
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 419 GCAGACCTCATCTCTCTTACCACCCGCTTCTTCCAGCCCATGAGCGCATTAACCTGG 478
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 AlaAsnLeuIleIleSerHisIleProIlePheThrAsnGlnAspSerAsnAspGluVal 81
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 479 AACACATGGAAGAG---CGCCTGTGATCCGGGCTGTGAGAACAGAGTGGTATCTAC 535
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 82 AsnTyrPheValAsnIleAspLeuIleGluLysIleLysAsnLysIleSerLeuIle 101
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 536 TCTCCTCATACAGCTATGATGCTGCGCCAGCGGCTCAACCAACTGGTGGCTAAAGG 595
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 102 HisLeuHisThrAlaPheAspAlaSerAlaAsnGlyMetSerMetGlnMetAlaLysArg 121
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 596 CTGGAGCTTGTACCTCCAGGCCCATACATCCTCCAAAGCTCCCACTACCTACAGAG 655
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 122 LeuGlyLeu----- 124
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 656 GGAACACACCGAGTAGAATTCAAGCTTAAGTACCAACCAAGCTGGACAAAGTCAATGCT 715
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 125 -----LeuAsnLeuLysGlnAspGluGlnAsnProTyrLeu 136
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 716 GCAGTGAAGAAATTGAC---GGTGTCTCTGTCACCTCTTTTCTGTAGAGCTGTAAT 772
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 137 ValValGlyGluLeuLysLeuGlyValSer----- 146
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 773 GAGAACAAACACGGATTAACTGAATTGTACTCAGAAAGCTTGTATGCAGGTGGTAGAT 832
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 147 -----ValAsp 148
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 833 TTTCTTTCCCGAACAACACTTTATCAGAAGACGGAATTTCTGCTACTGGAGAAGCT 892
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 149 TyrIleSerArg-----IleIleLysGlnLysPheLeuSer-----Pro 161
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 893 TTGCTTACATACTGATGAGAGGATGATGATGATGATGATGATGATGATGATGATGAT 952
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 162 IleIleLysTyrAsnAsnValPheArg----- 170
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 953 ACCATGATGTAGTGAATAAAAAGACACCTAAACATATCTCATATTGCTTACCCCTGGG 1012
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 170 ----- 170
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1013 GTGGGAGACCTTGAAGTCTCAAGTCAAGTGTGGCCCTGTGCTGTTCTGGG--- 1069
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 171 -----LeuGlnThrAsnLeuLysIleGlyIleIleGlyLysSerGlyTyr 186
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1070 -----AGCAGGTTCTGAGGGTGTGAGGCTGACCTTACCTCAGAGTGAGTANG 1120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

```

Db 187 LysPheAlaAspAspAlaPheAsnArgTyrGlnLeuAspMetLeuIleThrSerAspLeu 206
QY 1121 TCCCATCATGATGATTTGGATGCTGCTCCACAGGAATTAATGTCATCTCTGGAACAC 1180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 207 LysTyrHisAsnTyrPheAspAlaGlnAlaLysGlnAsnIleIleAspMetAsnHis 226
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1181 AGCAACACTGAACGAGCTTCTTCTGACCTTGAGATATAGCTG----- 1225
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 227 Leu---SerGlnSerIlePheIleAspValIleTyrAspGlnLeuThrLysPheTyrGly 245
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 1226 ---GATCTCAGCTTGAGAGATTAATTAATTAATC 1258
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 246 AsnAspAlaAsnLeuAsnLysSerLeuSerIleIle 257
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Search completed: August 22, 2003, 14:09:08  
 Job time : 75.5 secs